

ABSTRACT

The invention makes it possible to detect the characteristics of text data, and to analogize potential hidden meaning in the text data.

A word-cutting unit 3 performs a word-cutting process on the text data input from the input unit 1, a syntax-analysis unit 4 performs syntax analysis and a thesaurus-creation unit 5 creates thesauruses from the results, then after performing word cutting and syntax analysis again, a thesaurus-sorting unit 7 performs sorting, and a frequency-of-appearance unit calculates the frequency of appearance of the thesauruses, a correlation-coefficient-calculation unit 11 calculates correlation coefficients between thesauruses, a correlation-coefficient-total-calculation unit 13 for each thesaurus calculates the total of the correlation coefficients for each thesaurus, the graph-creation-display unit 15 creates a graph based on the frequency of appearance and total of the correlation coefficients for each thesaurus.

[Selected Drawing] Fig. 1